Internationalized Domain Names (IDNs)
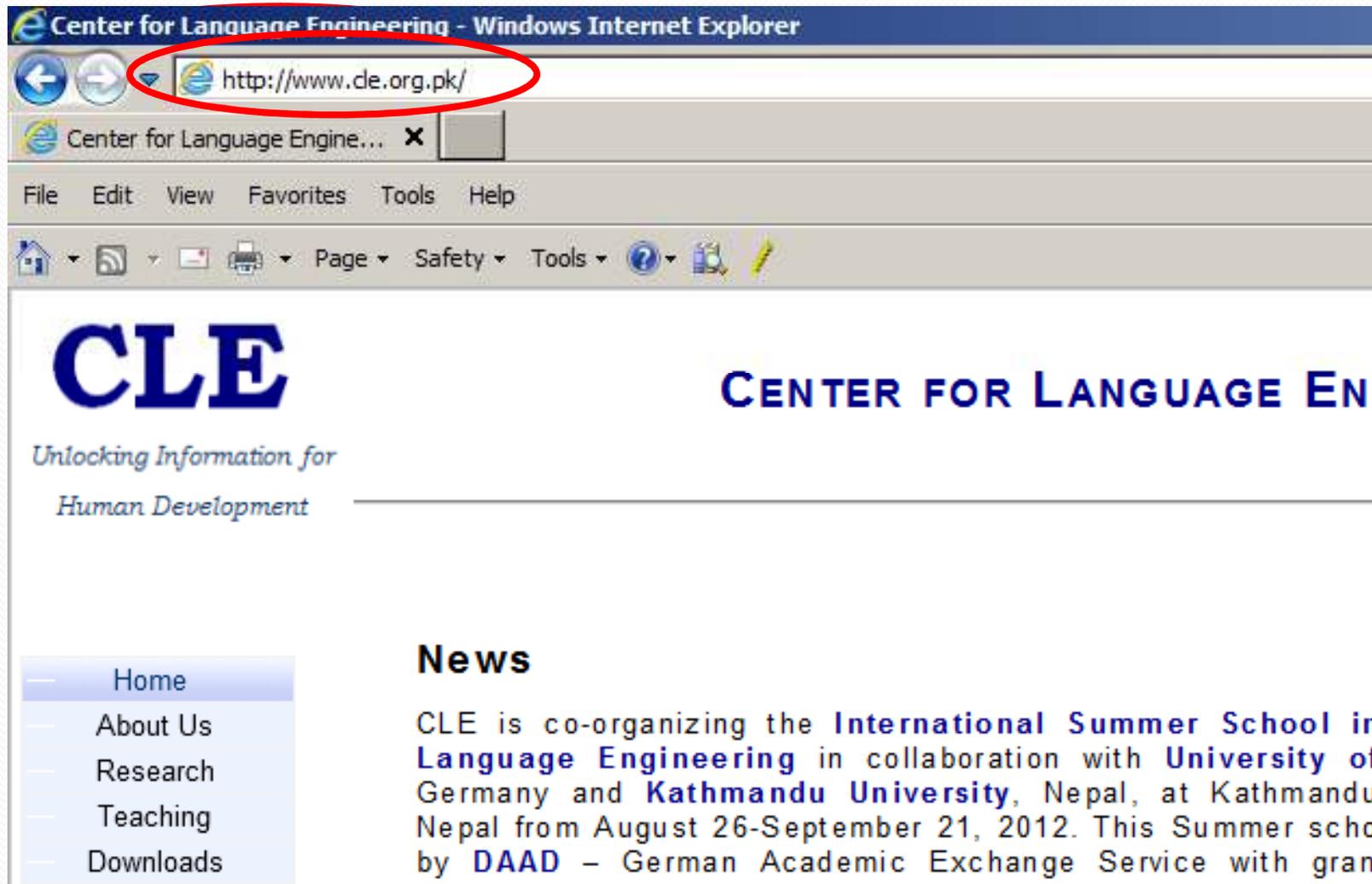
www.CLE.org.pk

DR. SARMAD HUSSAIN
PROFESSOR AND HEAD

Al- Khwarizmi Institute of Computer Sciences
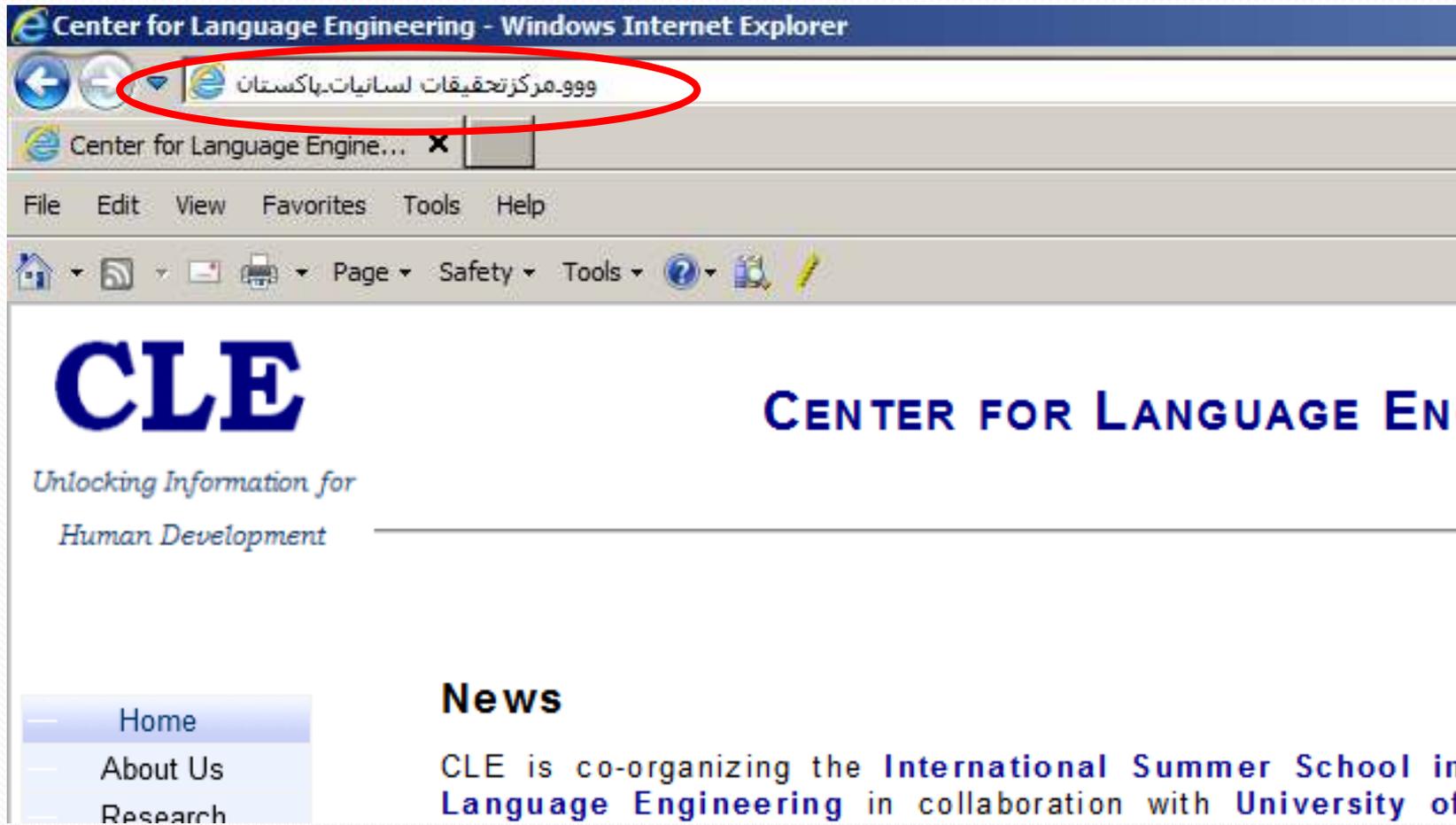**University of Engineering and Technology, Lahore**

sarmad.hussain@kics.edu.pk

Domain Name

Internationalized Domain Name

Background: Unicode

- Everything in the computers is represented as numbers
- Initially ASCII encoding:
  - A → 65
  - B → 66 …
- Only supported Latin script, primarily English
- Other encodings developed for other languages, but cumbersome to develop separate encoding for each language of the world

Unicode

- Thus effort started to develop Universal encoding or **UNIcode**
- Unicode Consortium develops the standard
- Covers all writing systems
- First version '*The Unicode Standard 1.0*' in 1991
- Current version '*The Unicode Standard 6.1*'

Unicode

- European scripts
  - Latin, Greek, Cyrillic, Armenian, Georgian, IPA
- Bidirectional (Middle Eastern) scripts
  - Hebrew, Arabic, Syriac, Thaana
- Indic (Indian and Southeast Asian) scripts
  - Devanagari, Bengali, Gurmukhi, Gujarati, Oriya, Tamil, Telugu, Kannada, Malayalam, Sinhala, Thai, Lao, Khmer, Myanmar, Tibetan, Philippine
- East Asian scripts
  - Chinese (Han) characters, Japanese (Hiragana and Katakana), Korean (Hangul), Yi

Unicode

- Other modern scripts
  - Mongolian, Ethiopic, Cherokee, Canadian Aboriginal
- Historical scripts
  - Runic, Ogham, Old Italic, Gothic, Deseret
- Punctuation and symbols
  - Numerals, math symbols, scientific symbols, arrows, blocks, geometric shapes, Braille, musical notation, etc.

Unicode is SCRIPT based

- One code per character per script
  - To avoid duplication of codes of same letter used by multiple scripts
  - For example:
    - The character code U+06A9 ک is same in Urdu, Sindhi, Pashto, Punjabi, Farsi, …
- Different code blocks reserved for different scripts
- For Arabic script 0600, 0601, …, 06FE, 06FF

Characters Semantics

- The Unicode standard includes an extensive database that specifies a large number of *character properties,* including:
  - Name
  - Type (e.g., letter, digit, punctuation mark)
  - Decomposition
  - Case and case mappings (for cased letters)
  - Numeric value (for digits and numerals)
  - Combining class (for combining characters)
  - Cursive joining behavior

Unicode

- Adopted by industry leaders as Apple, HP, IBM, Microsoft, etc.

- Supported in many platforms including Java, Linux and Microsoft Windows, etc.

- Supported by many internationalized applications including Open Office, Firefox, Thunderbird, Microsoft Office, etc.

Unicode is the basis for Internationalized Domain Names

- ✓ Background: Unicode
- Internationalized Domain Names (IDNs)
- Issues and challenges related to Arabic IDNs
- Sample (tentative solution) for Urdu language

Internationalized Domain Names (IDNs)

Domain Name System (DNS)

- Domain name is the address of a website which is used to access it

  e.g. www.cle.org.pk

Domain Name System (DNS)

www.cle.org.pk

1. www.cle.org.pk

6. Request Reply

4. 192.168.0.1

ISP

Host Server

5. **Requested Found / Not Found**

2. www.cle.org.pk

3. 192.168.0.1
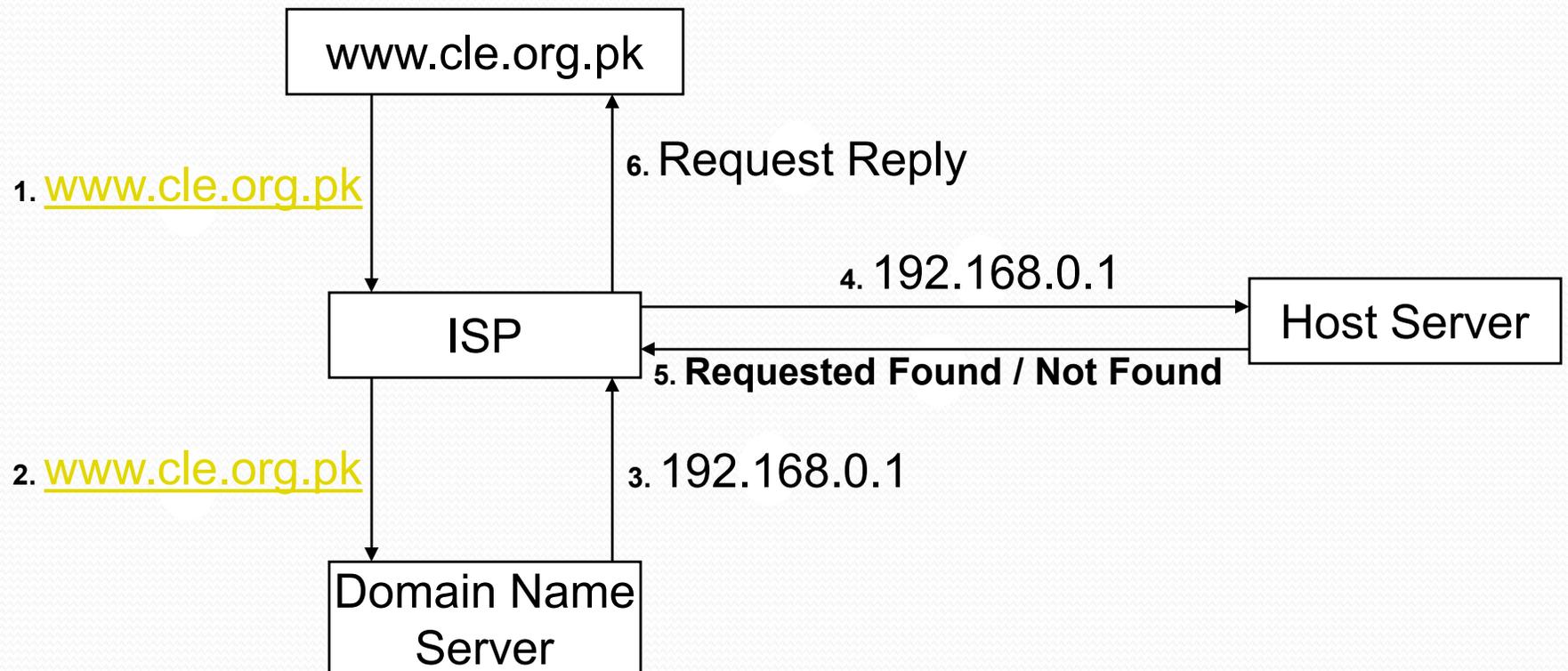
Domain Name Server

www.cle.org.pk = 192.168.0.1

Need of IDNs

- Domain name system (DNS) is in ASCII, i.e. Latin script
- Makes it difficult to access internet for people who do not understand English or Latin script

IDNs

- Basic reason for that is the internet addresses map into 7-bit ASCII standard
- We can not change the overall existing system
- The solution is to add layer that works on top of existing system
- IDN is any domain name consisting of labels which can be converted to ASCII format
- Initial set of protocols defined in 2003

IDNs

- A layer that takes the address in local languages and converts that into ASCII format
- DNS continues to resolve ASCII format addresses
- IDNs may be resolved at the User's computer

International Domain Name in Applications (IDNA)

```
┌─────────────────────────────────────────┐
│ www.مركزتحقيقات لسانيات .com              │
└─────────────────────────────────────────┘
        │                  ↑
        │   **1.** Convert to ASCII Compatible Encoding/punycode
        ↓                  │
┌─────────────────────────────────────────┐
│ http://www. xn--mgbaajbcq4ajo6na2af0a5bya11o.com/ │
└─────────────────────────────────────────┘
        │                  ↑
        │          **7.** Request Reply
        │                  │
        ↓                  │
               **5.** 192.16.0.1
┌──────────────┐  ─────────────────→  ┌──────────────┐
│     ISP      │                      │ Host Server  │
└──────────────┘  ←─────────────────  └──────────────┘
        │          ↑  **6. Requested Found / Not Found**
        │          │
        │       **4.** 192.16.0.1
        ↓          │
┌──────────────┐
│ Domain Name  │
│   Server     │
└──────────────┘
```
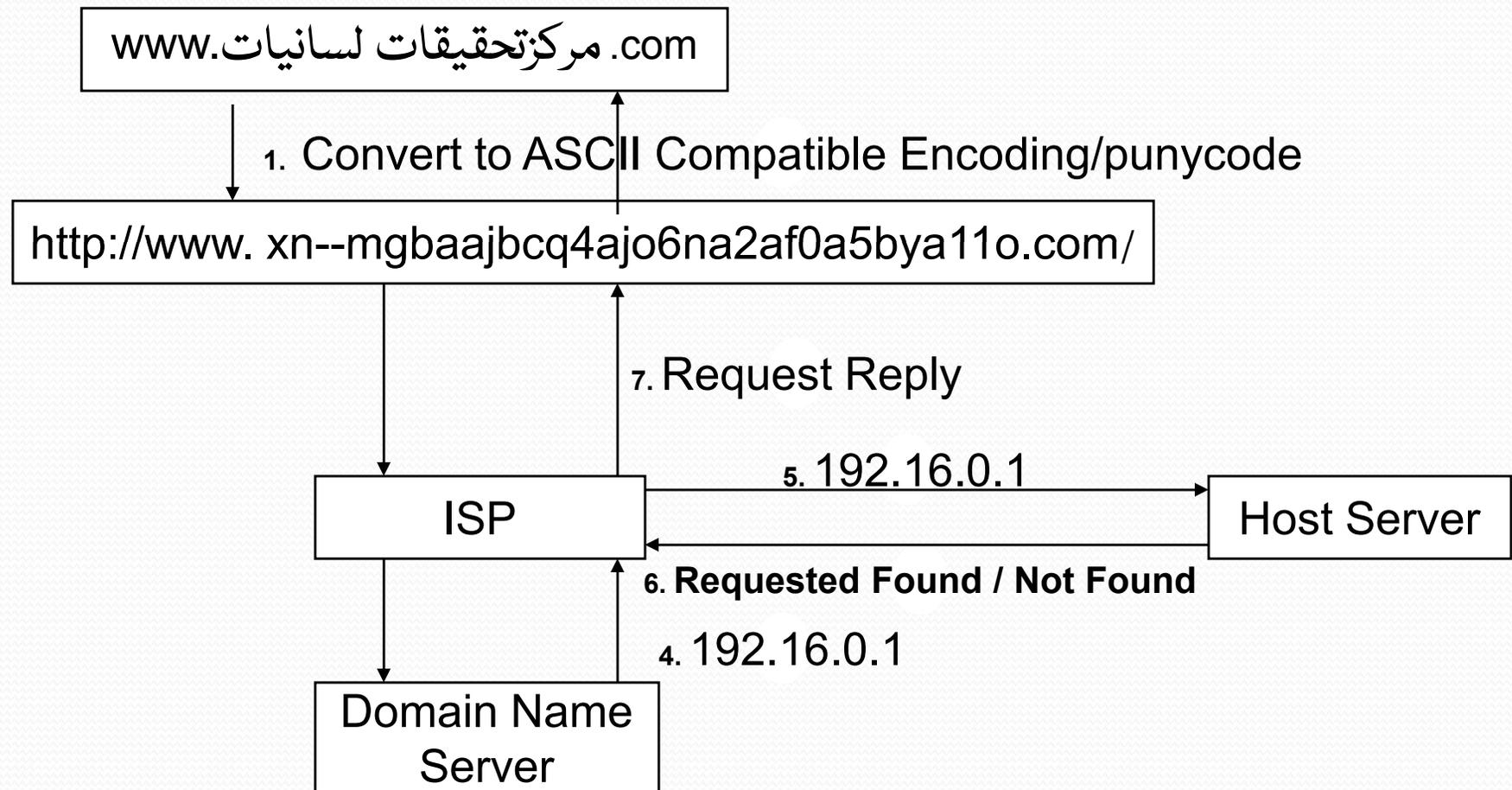
IDNA 2008

- Some Issues observed in the original IDNA2003 protocol
  - Exclusion based approach
  - Dependence on Unicode ver. 3.2
  - Hardcoded language specific separators
  - ZWNJ, ZWJ mapped onto null
  - Context independent
  - ...
- Decision to revise the original standard taken in 2006
- IDNA 2008 developed
  - RFCs 5890-5895

IDNA 2008

- Assigns values to all Unicode Character Database (UCD) on the basis of Unicode properties
  - The categories used are:
    - Ll - Lowercase_Letter
    - Lu - Uppercase_Letter
    - Lo - Other_Letter
    - Nd - Decimal_Number
    - Lm - Modifier_Letter
    - Mn - Nonspacing_Mark
    - Mc - Spacing_Mark

IDNA 2008

- Assigns values to all Unicode Character Database (UCD) on the basis of Unicode properties
  - Protocol-VALID (PVALID or allowed)
    - Registries should decide which subset to use based on relevant, confusion, etc.
  - CONTEXTJ or CONTEXTO
    - Joiners and Other characters requiring a contextual rule
  - DISALLOWED
    - Normalization
    - Case folding
    - Symbol or punctuation, etc. not used as part of a letter or digit
  - UNASSIGNED

IDNA 2008

- Exception list
  - PVALID -- Would otherwise have been DISALLOWED
    - 06FD; PVALID   # ARABIC SIGN SINDHI AMPERSAND
    - 06FE; PVALID   # ARABIC SIGN SINDHI POSTPOSITION MEN
  - CONTEXTO -- Would otherwise have been PVALID
    - 0660; CONTEXTO   # ARABIC-INDIC DIGIT ZERO ...
    - 0669; CONTEXTO   # ARABIC-INDIC DIGIT NINE
    - 06F0; CONTEXTO   # EXTENDED ARABIC-INDIC DIGIT ZERO ...
    - 06F9; CONTEXTO   # EXTENDED ARABIC-INDIC DIGIT NINE
  - DISALLOWED -- Would otherwise have been PVALID
    - 0640; DISALLOWED # ARABIC TATWEEL

Arabic IDN Variants

www.CLE.org.pk

DR. SARMAD HUSSAIN

PROFESSOR AND HEAD

Al- Khwarizmi Institute of Computer Sciences
University of Engineering and Technology, Lahore

Arabic Script

The unlabeled scripts of India are:
(west) Gurmukhi, Gujarati, Kannada, Malayalam,
and (east) Tamil, Telugu, Oriya, Bengali, Burmese.

Legend:
- Latin
- Cyrillic
- Hangeul (featural)
- Other alphabets
- Arabic
- Other abjads
- Devanagari
- Other abugidas
- Syllabaries
- Logographies

http://en.wikipedia.org/wiki/File:WritingSystemsoftheWorld4.png
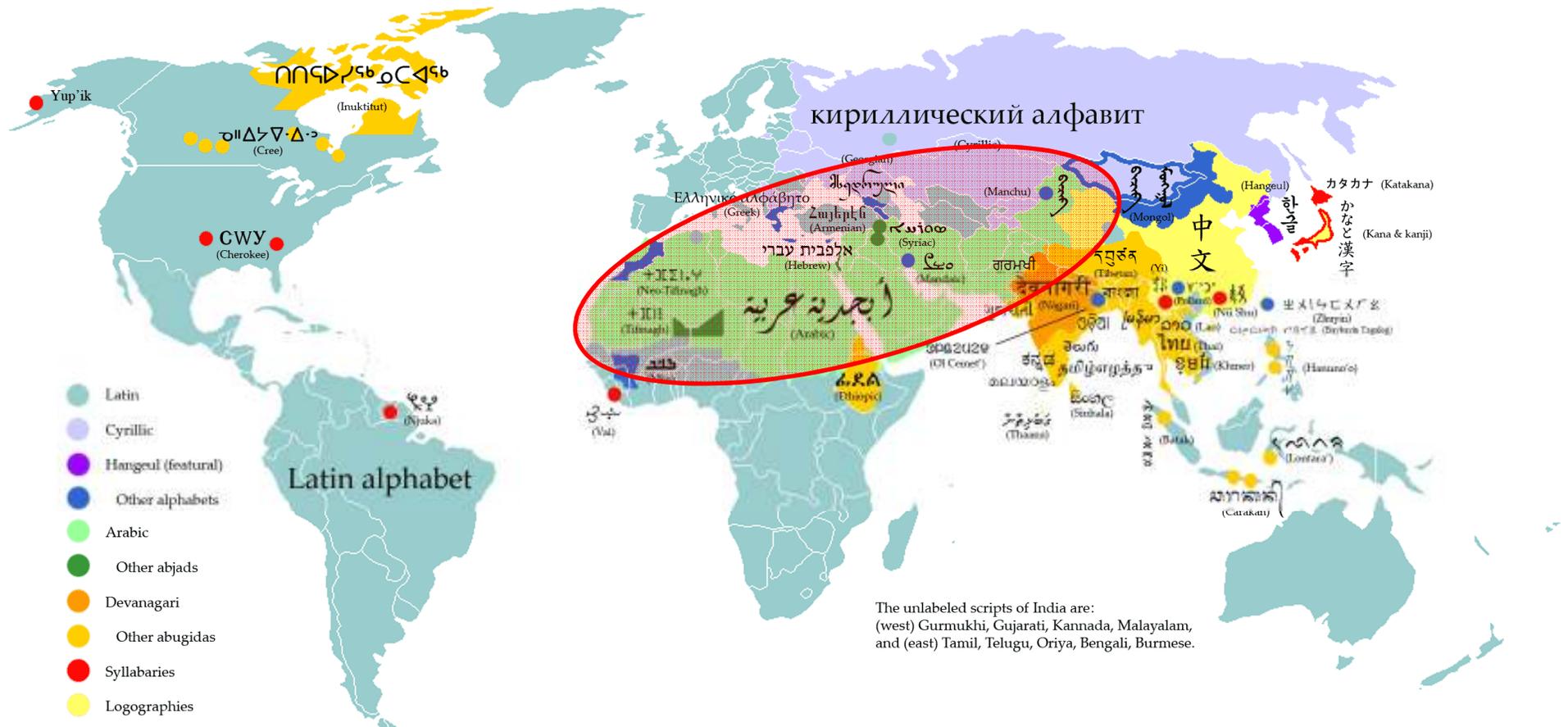
Arabic Script

- Writing system extended to represent multiple languages spoken in:

  - Middle East: *Arabic, Kurdish, Azerbaijani*
  - Africa: *Arabic, Bedawi, Huasa, ...*
  - Central Asia: *Kazakh, Uighur, Kirghiz, Azerbaijani*
  - South Asia: *Urdu, Pashto, Balochi, Sindhi, Kashmiri, Torwali, Burshuski, ...*
  - South East Asia: *Jawi*

Arabic Script

- Multiple writing styles
  - Naskh – Arabic, Sindhi, etc.
  - Nastalique – Persian, Urdu, Pashto, etc.
  - Others used frequently, but as stylistic variations
    - Kufi
    - Thuluth
    - Diwani
    - Riqa

وَ سخرَالشَمسَ وَالقمرَ
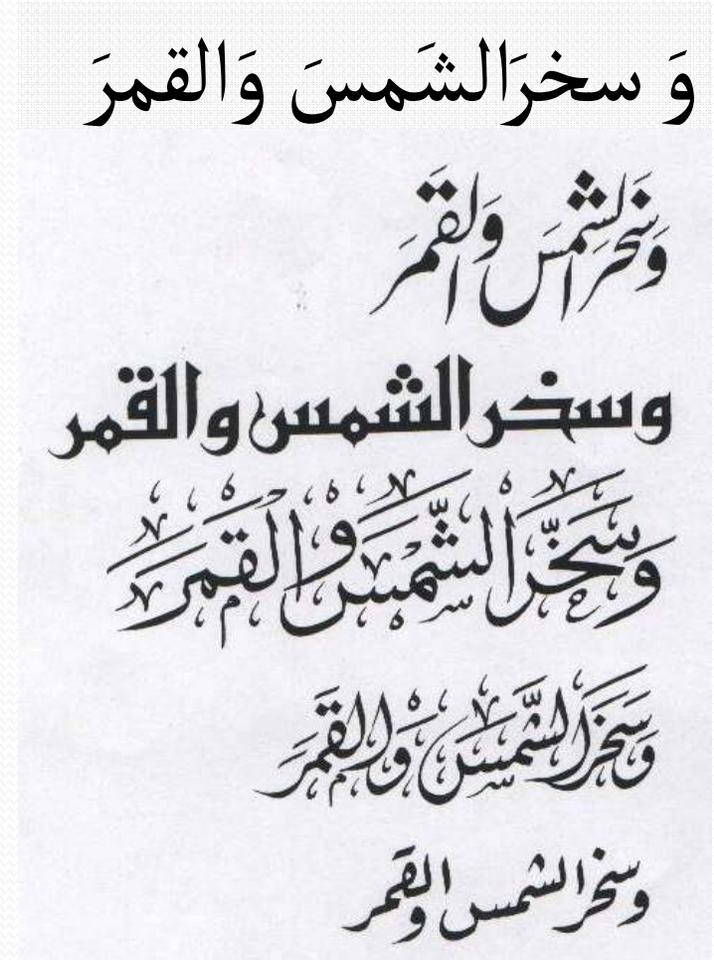
Arabic Script

- Consonantal (abjad)
  - Consonants written explicitly
  - Short vowels represented by optional vowel marks
  - Long vowels are represented by optional short vowel marks plus one of the three consonantal letters: ا و ى
- Bidirectional
  - Letters written from right to left
  - Digits written from left to right

ابجد ١٢١٢١١٣

→ ←

Arabic Script

- Cursive
  - Characters join with others before and after them
  - Four shapes: initial, medial, final and isolated for most characters

ب ببب چ چچچ س سسس

  - Some characters do no join with characters after them

ا ر ڑ ز ژ د ڈ ذ و ...
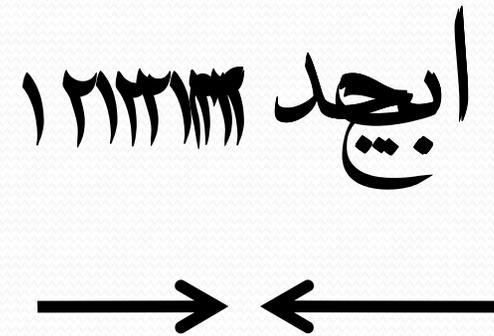
Positional Shaping

Isolated	Initial	Medial	Final
ب	با	کبا	کب
چ	چا	کچا	کچ
و	NA	NA	کو

Arabic Script

- Arabic script is defined from:
  - U+0600 to U+06FF
  - U+0750 to U+077F
  - U+FB50 to U+FDFF (Obsolete presentation forms)
  - U+FE70 to U+FEFF (Obsolete presentation forms except U+FDFx sequence)

Arabic Script IDNs

- The following are CONTEXTO by IDNA2008 but are not recommended

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	0600	ARABIC NUMBER SIGN	CONTEXTO	NO
	0601	ARABIC SIGN SANAH	CONTEXTO	NO
	0602	ARABIC FOOTNOTE MARKER	CONTEXTO	NO
	0603	ARABIC SIGN SAFHA	CONTEXTO	NO
	06DD	END OF AYAH	CONTEXTO	NO

Arabic Script IDNs

- The following are PVALID by IDNA2008 but these are not recommended

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
◌ؕ	0615	**ARABIC SMALL HIGH TAH**	**PVALID**	NO
ـ	0640	**ARABIC TATWEEL**	**PVALID**	NO
◌ٗ	0657	**ARABIC INVERTED DAMMA**	**PVALID**	NO
◌ٙ	0659	**ARABIC ZWARAKAY**	**PVALID**	NO
◌ٚ	065A	**ARABIC VOWEL SIGN SMALL V ABOVE**	**PVALID**	NO

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	065B	ARABIC VOWEL SIGN INVERTED SMALL V ABOVE	PVALID	NO
	065C	ARABIC VOWEL SIGN DOT BELOW	PVALID	NO
	065D	ARABIC REVERSED DAMMA	PVALID	NO
	065E	ARABIC FATHA WITH TWO DOTS	PVALID	NO
أ	0671	ARABIC LETTER ALEF WASLA	PVALID	NO
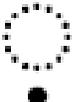
Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
ه	06D5	**ARABIC LETTER AE**	**PVALID**	NO
	06D6	**ARABIC SMALL HIGH LIGATURE SAD WITH LAM** **WITH**	**PVALID**	NO
	06D7	**ARABIC SMALL HIGH LIGATURE QAF WITH LAM** **WITH**	**PVALID**	NO
	06D8	**ARABIC SMALL HIGH MEEM** **INITIAL FORM**	**PVALID**	NO
	06D9	**ARABIC SMALL** **HIGH LAM ALEF**	**PVALID**	NO

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
ج	06DA	ARABIC SMALL HIGH JEEM	PVALID	NO
	06DB	ARABIC SMALL HIGH THREE DOTS	PVALID	NO
س	06DC	ARABIC SMALL HIGH SEEN	PVALID	NO
	06DF	ARABIC SMALL HIGH ROUNDED ZERO	PVALID	NO
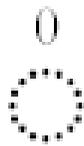
Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	06E0	ARABIC SMALL HIGH UPRIGHT RECTANGULAR ZERO	PVALID	NO
	06E1	ARABIC SMALL HIGH DOTLESS HEAD OF KHAH	PVALID	NO
	06E2	ARABIC SMALL HIGH MEEM ISOLATED FORM	PVALID	NO
	06E3	ARABIC SMALL LOW SEEN	PVALID	NO
	06E4	ARABIC SMALL HIGH MADDA	PVALID	NO

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	06E5	ARABIC SMALL WAW	PVALID	NO
	06E6	ARABIC SMALL YEH	PVALID	NO
	06E7	ARABIC SMALL HIGH YEH	PVALID	NO
	06E8	ARABIC SMALL HIGH NOON	PVALID	NO
	06EA	ARABIC EMPTY CENTRE LOW STOP	PVALID	NO

Arabic Script IDNs

- Following characters are DISALLOWED by IDNA 2008 and are not required

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
؋	060B	AFGHANI SIGN	DISALLOWED	N0
،	060C	ARABIC COMMA	DISALLOWED	N0
؍	060D	ARABIC DATE SEPARATOR	DISALLOWED	N0
؎	060E	ARABIC POETIC VERSE SIGN	DISALLOWED	N0
؏	060F	ARABIC SIGN MISRA	DISALLOWED	N0

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	061B	ARABIC SEMICOLON	DISALLOWED	N0
	061E	ARABIC TRIPLE DOT PUNCTUATION MARK	DISALLOWED	N0
	061F	ARABIC QUESTION MARK	DISALLOWED	N0
	066A	ARABIC PERCENT SIGN	DISALLOWED	N0
	066B	ARABIC DECIMAL SEPARATOR	DISALLOWED	N0

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
,	066C	ARABIC THOUSANDS SEPARATOR	DISALLOWED	N0
★	066D	ARABIC FIVE POINTED STAR	DISALLOWED	N0
ٵ	0675	ARABIC LETTER HIGH HAMZA ALEF	DISALLOWED	N0
ٶ	0676	ARABIC LETTER HIGH HAMZA WAW	DISALLOWED	N0
ٷ	0677	ARABIC LETTER U WITH HAMZA ABOVE	DISALLOWED	N0
ٸ	0678	ARABIC LETTER HIGH HAMZA YEH	DISALLOWED	N0

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	06D4	**ARABIC LETTER HAH WITH HAMZA ABOVE**	**DISALLOWED**	N0
	06DE	**ARABIC START OF RUB EL HIZB**	**DISALLOWED**	N0
	06E9	**ARABIC PLACE OF SAJDAH**	**DISALLOWED**	N0

Arabic Script IDNs

- Following are PVALID characters by IDNA2008 decision and are required

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
	0610	ARABIC SIGN SALLALLAHOU ALAYHE WASSALLAM	PVALID	YES
	0611	ARABIC SIGN ALAYHE ASSALLAM	PVALID	YES
	0612	ARABIC SIGN RAHMATULLAH ALAYHE	PVALID	YES
	0613	ARABIC SIGN RADI ALLAHOU ANHU	PVALID	YES
	0614	ARABIC SIGN TAKHALLUS	PVALID	YES

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
ء	0621	ARABIC LETTER HAMZA	PVALID	YES
آ	0622	ARABIC LETTER ALEF WITH MADDA ABOVE	PVALID	YES
أ	0623	ARABIC LETTER ALEF WITH HAMZA ABOVE	PVALID	YES
ؤ	0624	ARABIC LETTER WAW WITH HAMZA ABOVE	PVALID	YES
إ	0625	ARABIC LETTER ALEF WITH HAMZA BELOW	PVALID	YES (Variant of Base Character)

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
ئ	0626	**ARABIC LETTER YEH WITH HAMZA ABOVE**	PVALID	YES (Variant of Base Character)
ا	0627	**ARABIC LETTER ALEF**	PVALID	YES
ب	0628	**ARABIC LETTER BEH**	PVALID	YES
ة	0629	**ARABIC LETTER TEH MARBUTA**	PVALID	YES (Variant of Base Character)
ت	062A	**ARABIC LETTER TEH**	PVALID	YES
ث	062B	**ARABIC LETTER THEH**	PVALID	YES

Character	Unicode	Description	Current status in IDNA 200x	Recommen-dation
ج	062C	ARABIC LETTER JEEM	PVALID	YES
ح	062D	ARABIC LETTER HAH	PVALID	YES
خ	062E	ARABIC LETTER KHAH	PVALID	YES
د	062F	ARABIC LETTER DAL	PVALID	YES
ذ	0630	ARABIC LETTER THAL	PVALID	YES

Sources of Variants

- intrinsically, some strings are considered equivalent by a language or script community

- extrinsically, script encoding scheme (Unicode) is non-optimal, introducing additional ambiguity for end-users

Motivation

- Without Variant Management
  - security threats, allowing for easy phishing
  - perception of a broken internet experience by end-user

Issues

- Technical
  - Required combining marks
  - Optional combining marks
  - Same shape in a particular position
  - Similar shape in a particular position
  - Digits
  - Joining characters – ZWNJ
- User Interface
  - Input method variation
  - Bidirectional rendering issues in applications
  - General rendering issues in applications
- Policy
  - Specification of language table
  - Bundling, blocking, reserving

Required Combining Marks

Combining Mark	Composed Form	Decomposed Form	Unicode Normalized Form
◌ٓ U+0653	آ U+0622	آ U+0627 U+0653	Defined
◌ؕ U+0615	ڑ U+0691	رؕ U+0631 U+0615	Not Defined
◌ۛ U+065B	◌ۛ U+06EE	دۛ U+062F	Not Defined

Can you guess: ڑ ڑ

Optional Combining Marks

- Vowel Marks
- Honorifics
- Consonantal Gemination
- ...

أَحد = احد

أُحد = احد

أُحد ≠ أَحد

اٰ U+0627 U+0653	اٰ	ا U+0627	**Not same**
اُ U+0627 U+64F	اُ	ا U+0627	**same**

55

Same Shape

Unicode	Initial Form	Medial Form	Final Form	Isolated Form
ک U+06A9	ک	کـ	ک	ک
ك U+0643	ک	کـ	لک	ك
ة U+0629	-	-	ة	ة
ة U+06C3	-	-	ـة	ة

Can you guess: پاکستان پاکستان

56

Similar Shape

Unicode	Initial Form	Medial Form	Final Form	Isolated Form
ک U+06A9	ک	ـکـ	ـک	ک
ڪ U+06AA	ڪ	ـڪـ	ـڪ	ڪ
ت U+062A	تـ	ـتـ	ـت	ت
ٺ U+067A	ٺـ	ـٺـ	ـٺ	ٺ

پاکستان پاکستان پاکستان پاکستان پاکستان

Digits

ASCII	Arabic-Indic	Eastern Arabic-Indic
(U+0030) 0	(U+0660)٠	(U+06F0)۰
(U+0031) 1	(U+0661)١	(U+06F1)۱
(U+0032) 2	(U+0662)٢	(U+06F2)۲
(U+0033) 3	(U+0663)٣	(U+06F3)۳
(U+0034) 4	(U+0664)٤	(U+06F4)۴
(U+0035) 5	(U+0665)٥	(U+06F5)۵
(U+0036) 6	(U+0666)٦	(U+06F6)۶
(U+0037) 7	(U+0667)٧	(U+06F7)۷
(U+0038) 8	(U+0668)٨	(U+06F8)۸
(U+0039) 9	(U+0669)٩	(U+06F9)۹

Different or Same?

123abc123
۱۲۳abc۱۲۳
۱۲۳abc۱۲۳
۱۲۳abc۱۲۳
۱۲۳abc۱۲۳
۱۲۳abc123
123abc۱۲۳
۱۲۳abc123
123abc۱۲۳

ZWNJ

- Exact variants ?
  - طب
  - ط ب
  - طب
  - ب ZWNJ ط
- Non-exact variants?
  - خوبصورت
  - خوبصورت
  - خوب-صورت

User Interface Issues

- Typing a string in
  - application V
  - for language W
  - in country X
  - with keyboard Y
  - in operating system Z

Application Interfaces

61

Application Interfaces

Policy

- Complete and correct "Language" table
  - Mechanisms to record characters
  - Mechanisms to record variants
  - Mechanisms to record rules
    - ZWNJ
    - Script Mixing
    - Digit Mixing
    - ...
  - Mechanisms to record differences in registration vs. resolution (e.g. kafs)
- Mechanisms to record registration differences
  - Variant bundling/blocking/reservation

- Need for a single language tables for all the languages of Pakistan
  - Balochi
  - Pashto
  - Punjabi
  - Seraiki
  - Sindhi
  - Urdu
  - Torwali
- Collective Issues for multiple languages

IDN ccTLD پاکستان

www.CLE.org.pk

Dr. Sarmad Hussain

Professor and Head

Al- Khwarizmi Institute of Computer Sciences
University of Engineering and Technology, Lahore

ICANN IDN ccTLD Fast Track Process

1

Preparation
— Build community consensus on:
- which IDN ccTLD to apply for
- how it will be run
- which organisation will operate it
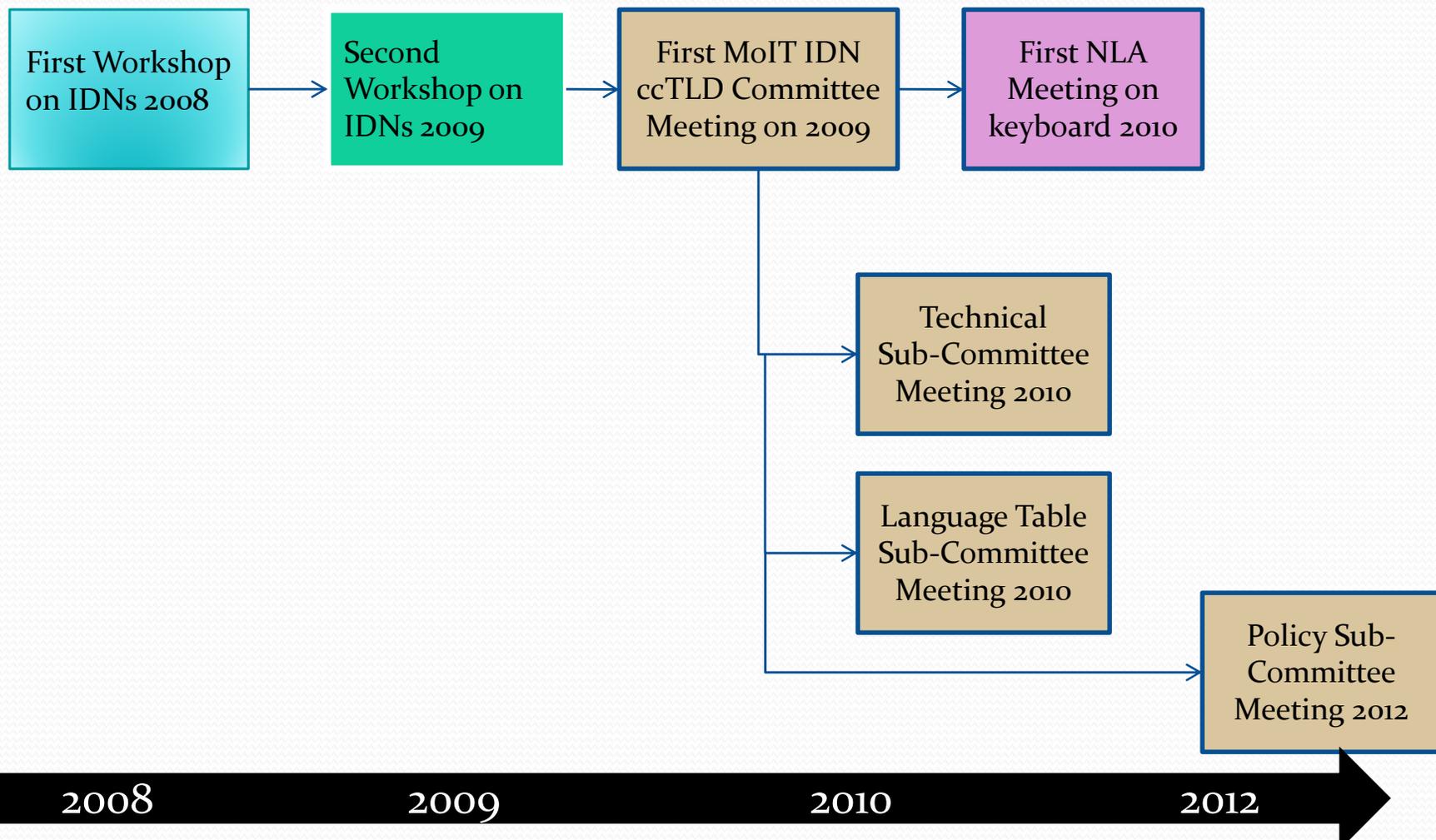— Prepare supporting documentation

2

String Evaluation
— Determines what string can be used to represent the country
— Evaluation of technical and linguistic string requirements
— Submit a request via the Online Request System

3

Delegation
— Determines which entity can manage the domain
— Follows the same process used for ASCII-based ccTLDs
— Submit an application to IANA root zone management

Sequence of Events

First Workshop on IDNs 2008 → Second Workshop on IDNs 2009 → First MoIT IDN ccTLD Committee Meeting on 2009 → First NLA Meeting on keyboard 2010

First MoIT IDN ccTLD Committee Meeting on 2009 → Technical Sub-Committee Meeting 2010

→ Language Table Sub-Committee Meeting 2010

→ Policy Sub-Committee Meeting 2012

2008   2009   2010   2012

- IDN ccTLD Preparation
  - String Decision (completed)
    - Second Workshop 2009
  - Language Table (completed)
    - Language Table Sub-Committee Meeting 2010
- IDN ccTLD Application and Evaluation (2010; accepted)
- IDN ccTLD Delegation
  - Technical Test bed (completed)
    - Technical Sub-Committee Meeting 2010
  - Policy Development (in progress)
    - Third Workshop on IDNs 2012
  - Manager Selection
  - Application for Delegation

First Workshop on IDNs 2008

- The first workshop on IDNs was held at NUCES, Lahore on 20th April, 2008
- Attended by MoIT, language authorities and academies, language, linguistics and technology experts from academia and other relevant organizations
- Need for IDNs was explained
- A number of language and technical experts collaborated to develop individual language tables

Outcomes

- First draft of language tables created:
  - Balochi
  - Pashto
  - Punjabi
  - Saraiki
  - Sindhi
  - Torwali
  - Urdu
- Can be accessed at

http://www.crulp.org/idn/download/LanguageTables.pdf

Second Workshop on IDNs 2009

- A follow-up workshop was organized by the National IT Development and Promotional Unit (NIDU) Ministry of IT on May 15-16, 2009 at NUCES, Lahore
- Divided in two sessions
  - Open public session on first day to seek public opinion on all IDN issues (invitation through newspaper advertisement)
  - Closed meeting of a group of technical and language experts on second day to finalize issues based on public feedback

Outcomes

- Major outcomes of the second workshop are as follows
  - Single language table to be adopted for all Pakistani languages
  - پاکستان ـ to be IDN ccTLD string for Pakistan
  - '.' to be used as label separator
  - ZWNJ and '-' (hyphen-minus) to be used as character separators
  - Extended Arabic Indic and ASCII digits to be used in Pakistani
  - IDNs where both cannot be mixed in a single label
  - Confusable or variant characters to be mapped onto each other during registration
- Complete minutes of the workshop can be accessed at

http://www.crulp.org/idn/IDN2009/techdoc.htm

IDN ccTLD Committee Meeting 2009

- The Ministry of IT formed a committee for the implementation of پاکستان۔ ccTLD registry
- Members include representatives from the following
  - Pakistan Software Export Board, Islamabad
  - E-Government Directorate, Islamabad
  - National Language Authority, Islamabad
  - Pakistan Computer Bureau, Islamabad
  - Computer Society of Pakistan, Islamabad
  - Pakistan Software House Association, Islamabad
  - ISP Association of Pakistan, Islamabad
  - Academia
  - Civil Society
- First meeting convened on 22nd October, 2009 under the chairmanship of Joint Secretary (Development), MoIT

Outcomes

- Major outcomes of the meeting were the following
  - Minutes of the second workshop on IDNs held on May 15-16, 2009 at NUCES, Lahore were unanimously approved and adopted by the Technical Committee for implementation of پاکستان. ccTLD registry
  - Three separate sub-committees established
    1. Technical Sub-Committee (to resolve technical issues)
    2. Language Table Sub-Committee (to finalize language tables)
    3. Policy Sub-Committee (to develop policies)
- NLA to work for including '.' on Urdu keyboard

NLA Meeting on Keyboard 2010

- National Language Authority meeting held in 2010
- The purpose of the meeting was to review Urdu keyboard design in the context of IDNs
- The committee decided the following
  - Have multiple versions of KB based on major languages
  - Each version support all characters in language table
  - Additional characters for email, Internet, etc. to be included
  - To be done with least change in existing KB version for backward
  - compatibility
- Final keyboard design after language table for Pakistani IDNs is finalized
- Complete detailed minutes to be released by NLA

Technical Sub-Committee Meeting 2010

- The first technical sub-committee meeting was held on 25th March, 2010 at Pakistan Software Export Board, Islamabad
- The following agenda items were discussed
  - Progress on setting up test-bed for پاکستان۔ IDN ccTLD
  - Requirements for setting up test-bed registry for پاکستان۔ IDN ccTLD
  - Discussion of future steps needed for the implementation of پاکستان۔ IDN ccTLD, components and completion time frame
- Outcomes
  - Technical issues relating to test-bed operation resolved
- Complete detailed minutes to be released by PSEB

Language Table Sub-Committee Meeting 2010

- This meeting was held on March 27, 2010 at National University of Computer and Emerging Sciences, Lahore
- The purpose of this meeting was to finalize the language table for Pakistani languages for پاکستان. IDN ccTLD
- Outcomes of the meeting were
  - Single language table listing complete set of characters required to represent all languages of Pakistan
  - Normalization table for characters in language table
  - Extended normalization table for characters in language table
  - Variant table corresponding to the language table

Third Workshop on IDNs 2012

- Two days workshop was organized by the Ministry of IT on April 20-21, 2012 at Center for Language Engineering (CLE), Al-Khawarizmi Institute of Computer Science (KICS), UET Lahore.

- First day, public meeting advertised through print and online media

- Second day, closed meeting of policy sub-committee for policy finalization for پاکستان۔ IDN ccTLD

Outcomes

- Policy recommendations for the following issues were finalized in this meeting
  - Translations for reserved second levels labels
  - Registration of domain names at second level
  - Pricing policy for second and third level domain names
  - Sunrise period for پاکستان. IDN ccTLD
  - Reserved words for پاکستان IDN ccTLD
  - Single letter domain names reservation for پاکستان IDN ccTLD
  - Domain limit for پاکستان IDN ccTLD
  - Variants limit for پاکستان IDN ccTLD
  - Pricing policy regarding variants
  - Domain name registration data (DNRD)
- Minutes of workshop will be published after the approval from MoIT at

http://www.cle.org.pk/IDN/idn2012/

Outcomes

- Policy recommendations for the following issues were finalized in this meeting
  - Mode of DNRD submission
  - Language of DNRD
  - Registrant eligibility for پاکستان IDN ccTLD
  - Governance model
  - Industry model for پاکستان IDN ccTLD
  - Cost of domain name
  - Highly demanded domain names
  - Dispute resolution process for پاکستان IDN ccTLD
  - Privacy services for پاکستان IDN ccTLD
  - Evaluation of پاکستان IDN ccTLD services
- Minutes of workshop will be published after the approval from MoIT at

http://www.cle.org.pk/IDN/idn2012/

Delegation Process

- Delegation to an organization which is supported through a community process
  - To ensure continued community involvement

- Delegation organization must be technically competent
  - to ensure security and stability of the organization

Administrative and Technical skills of the proposed operator

- Organization's structure, key Personnel
  - Financial officers
  - Business officers
  - Overview of staff
- Documentation of administrative and technical contacts
  - Identities
  - Contact details
  - roles in organization
- Explanation of Internet management and registry operations experience
- Technical capabilities of the sponsoring organization
  - Technical plan for registry
  - DNS operations
  - Physical configuration of the registry
  - technical facilities

Administrative and Technical skills of the proposed operator(Contd)

- Governing policies
  - How these policies are settled
  - Evolution framework
- Information on security
  - Database protection
  - Disaster recovery
  - Hacking attacks
- Scaling of registry operations with time
- Domain registration model
  - registry-registrar model
  - Direct interaction with the registry
- Mechanism of including Local Internet community feedback

Administrative and Technical skills of the proposed operator(Contd)

- Contractor or Third party Role
  - Stability issue handling
  - Disaster Recovery
- In case of distributed operation to other parties, how the sponsoring organization will maintain responsibility?

Administrative and Technical skills of the proposed operator(Contd)

- Internet connectivity of the organization, and the services it will provide
  - redundancy
  - resiliency provisions
  - provisions for IPv6 connectivity
- description of the configuration and plan for the name server constellation that will support name resolution for the top-level domain
- registry's interfaces to the community
- information on zone data generation, and the provision of public WHOIS services
- information on resiliency
  - system outages handling
  - system recovery
  - escrow procedures in the event of a significant disasters

Administrative and Technical skills of the proposed operator(Contd)

- Description of technology that will be used
  - Support of IDNs?
  - DNSSEC implementation?
- processes and plans to ensure operations remain in line with global standards
  - Best practices
  - security and stability concerns
- how all these relevant skills and operations will be appropriately managed in country
- timelines and deployment strategies relating to deployment of the registry technical platform and staffing of the registry, if the registry is not already fully commissioned

Organization's Legal Status

- ccTLDs assigned to individuals in past
- Now ccTLDs are assigned to Organizations
- Legal authenticity, status, character of organization
- Detail information about organization
  - Full name, Address, telephone, fax, registration proof
- Additional information
  - History of organization
  - Directors, officers, managers
- Information will be helpful in evaluation

Chronology of community consultation events, letters from representatives of the local internet community

- Interests of the Internet user community
- Feedback and consensus about ccTLD from community
- Consensus process must be described by the applicant
- Statements from significant community representatives can be tendered like
  - User groups
  - Internet Organizations
  - ISPs
  - Trade groups
- Documentation of opposition to the proposal
- Documentation for local community support
- Summary of the intended administrative operation of the domain name

Chronology of community consultation events letters from representatives of the local internet community (contd)

- Participation of community in management of domain
- Applicant should explain the structure of sponsoring organization
- Selection criteria of the management
- How community input is absorbed in evolution of organization's policy and operations?

Government contact

- ccTLD is an important resource of country's internet community
- Government support must be indicated in application (for example IT ministry).
- Applicant should explain the views of the government
- Applicant must indicate the government laws related to ccTLD administration
- Support documents from government authorities

Fair and Equitable Treatment

- Unbiased functionality
- Domain registration policy be provided to IANA
- Users of the domain are not required to use a certain Internet service provider, or use a particular type of software
- Limitation on registration process may be applied on geographical basis i.e. only nationality holder can apply for domain

Original documentation

- Originals, or certified copies, of all official documents relating to the request must be sent to IANA Root Zone Management team
  - **IANA Root Management
    Ticket Number:
    4676 Admiralty Way, Suite 330
    Marina del Rey, CA 90292
    USA**
- Reference number must be indicated which is provided at the time of template submission
- All documentation must be sent in English language to avoid translation errors thus avoiding delays in the delegation process

Thanks

sarmad.hussain@kics.edu.pk